Journal of Computational and Applied Mathematics 236 (2012) 3445–3460

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Tail asymptotics of the queue size distribution in the M/M/m retrial queue[☆]

Jerim Kim[a], Jeongsim Kim[b], Bara Kim[a,*]

[a] Department of Mathematics, Korea University, 145, Anam-ro, Seongbuk-gu, Seoul, 136-701, Republic of Korea

[b] Department of Mathematics Education, Chungbuk National University, 52 Naesudong-ro, Heungdeok-gu, Cheongju, Chungbuk, 361-763, Republic of Korea

ARTICLE INFO

Article history:
Received 20 April 2010
Received in revised form 31 January 2012

Keywords:
M/M/m retrial queue
Queue size distribution
Censored Markov process
Tail asymptotics
Karamata Tauberian theorem
Riemann–Lebesgue lemma

ABSTRACT

We consider an M/M/m retrial queue and investigate the tail asymptotics for the joint distribution of the queue size and the number of busy servers in the steady state. The stationary queue size distribution with the number of busy servers being fixed is asymptotically given by a geometric function multiplied by a power function. The decay rate of the geometric function is the offered load and independent of the number of busy servers, whereas the exponent of the power function depends on the number of busy servers. Numerical examples are presented to illustrate the result.

1. Introduction

Retrial queues are queueing systems in which arriving customers who find all servers occupied may retry for service again after a random amount of time. Retrial queues have been widely used to model many problems in telephone systems, call centers, telecommunication networks, computer networks and computer systems, and in daily life. Detailed overviews for retrial queues can be found in the bibliographies [1–3], the surveys [4–6], and the books [7,8].

In this paper we consider an M/M/m retrial queue where customers arrive from outside the system according to a Poisson process with rate λ. The service facility consists of m identical servers, and service times are exponentially distributed with mean μ^{-1}. If there is a free server when a customer arrives from outside the system, this customer begins to be served immediately and leaves the system after the service is completed. On the other hand, any customer who finds all the servers busy upon arrival joins a retrial group, called an orbit, and then attempts service after a random amount of time. If there is a free server when a customer from the orbit attempts service, this customer receives service immediately and leaves the system after the service completion. Otherwise the customer comes back to the orbit immediately and repeats the retrial process. The retrial time, i.e., the length of the time interval between two consecutive attempts made by a customer in the orbit, is exponentially distributed with mean ν^{-1}. The arrival process, the service times, and the retrial times are assumed to be mutually independent. The traffic load ρ is defined as $\rho = \frac{\lambda}{m\mu}$. It is known that the M/M/m retrial queue is stable if and only if $\rho < 1$ [9]. We assume that $\rho < 1$ for stability of the system.

The M/M/m retrial queue has been studied by several authors. Greenberg and Wolff [10] studied an approximation method for steady-state probability and provided an upper bound on system performance for the M/M/m retrial queue

☆ The second author's research was supported by the Korea Research Foundation (KRF) grant funded by the Korea government (MEST) (2009-0076674). The third author's research was supported by the Korea Research Foundation (KRF) grant funded by the Korea government (MEST) (2009-0076600).

* Corresponding author.
E-mail addresses: b1155@hanmail.net (J. Kim), jeongsimkim@chungbuk.ac.kr (J. Kim), bara@korea.ac.kr (B. Kim).

0377-0427/$ – see front matter
doi:10.1016/j.cam.2012.03.027

with a finite capacity. Pearce [11] found the joint distribution of the number of customers in the orbit and the number of busy servers for the M/M/m retrial queue with the orbit of finite capacity. For the M/M/m retrial queue, Neuts and Rao [12] presented a simplifying approximation of the joint distribution of the number of customers in the orbit and the number of busy servers at steady state by placing a fictitious limit in the orbit capacity. This approximation based on truncation method, starting from the pioneering paper in [13], has been widely used in the numerical analysis of retrial queues.

Analytical studies of the multiserver retrial models are limited and in many cases restricted to two-server systems. Keilson et al. [14] established a recursive algorithm for the computation of steady-state probabilities in the M/M/2 retrial queue. Hanschke [15] showed that the generating functions of the steady-state probabilities can be expressed in terms of generalized hypergeometric functions in the M/M/2 retrial queue.

The main contribution of our work is that we find the tail asymptotics for the distribution of the queue size (i.e., the number of customers in the orbit) in the M/M/m retrial queue. Tail behaviors of the queue size and the waiting time distributions in retrial queues began to be investigated recently. Nobel and Tijms [16] and Kim et al. [17] studied light-tailed asymptotic behaviors in the M/G/1 retrial queue where the service time distribution has a finite exponential moment. Nobel and Tijms [16] suggested a light-tailed approximation of the waiting time distribution. Kim et al. [17] showed that the tail of the queue size distribution is asymptotically given by a geometric function multiplied by a power function. Kim et al. [18] extended the result of [17] to the case of MAP/G/1 retial queue. On the other hand, Shang et al. [19] and Kim and Kim [20] studied heavy-tailed asymptotics in the M/G/1 retrial queue. Shang et al. [19] showed that the stationary distribution of the queue size in the M/G/1 retrial queue is subexponential if the stationary distribution of the queue size in the corresponding ordinary M/G/1 queue is subexponential. As a corollary of this property, they proved that the stationary distribution of the queue size has a regularly varying tail if the service time distribution has a regularly varying tail. Kim and Kim [20] showed that if the service time distribution has a regularly varying tail of index $-\alpha$, $\alpha > 1$, then the waiting time distribution has a regularly varying tail of index $1-\alpha$.

In this paper we investigate the tail asymptotics for the joint distribution of the number of customers in the orbit and the number of busy servers at steady state in the M/M/m retrial queue. More precisely, we show that for $i = 0, 1, \dots, m$,

$$\mathbb{P}(N = n, S = i) \sim \frac{c}{i!}\left(\frac{\nu}{\mu}\right)^i n^{\frac{\lambda}{m\nu}-m+i}\rho^n \quad \text{as } n \to \infty,$$

where N is the number of customers in the orbit at steady state, S is the number of busy servers at steady state, and c is the positive constant. As shown in the above formula, the stationary queue size distribution with the number of busy servers being fixed, is asymptotically given by a geometric function multiplied by a power function. The decay rate of the geometric function is the offered load ρ and independent of the number of busy servers, whereas the exponent of the power function depends on the number of busy servers.

In order to derive the main result, we first consider a censored Markov process obtained by observing the M/M/m retrial queue only when the number of busy servers is less than or equal to $m-1$. A matrix differential equation is derived for the vector probability generating function of the stationary distribution of the censored Markov process. The result is obtained by studying analytic properties of the solution of the differential equation.

The remainder of this paper is organized as follows. In the next section, we briefly review our model and introduce notations. In Section 3, we present our main result on the tail asymptotics of the queue size distribution. Section 4 is devoted to the derivation of the tail asymptotics stated without proof in Section 3. In Section 5, numerical examples are presented to illustrate the result. In Appendix, we prove the results stated without proof in Section 4.

2. The M/M/m retrial queue

We consider the M/M/m retrial queue as described in Section 1. By a little abuse of notation, let $N(t)$ denote the number of customers in the orbit at time t and $S(t)$ the number of busy servers at time t. Then $\{(N(t), S(t)) : t \geq 0\}$ is a continuous time Markov process with state space $\{0, 1, 2, \dots\} \times \{0, 1, \dots, m\}$. The infinitesimal generator Q of $\{(N(t), S(t)) : t \geq 0\}$ is given by

$$Q = \begin{bmatrix} B_0 & A_0 & O & O & \cdots \\ C_1 & B_1 & A_1 & O & \cdots \\ O & C_2 & B_2 & A_2 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix},$$

where $A_n, B_n,\ n \geq 0$ and $C_n,\ n \geq 1$, are $(m+1) \times (m+1)$ matrices whose (i,j) components $(A_n)_{ij}$, $(B_n)_{ij}$ and $(C_n)_{ij}$, $0 \leq i,\ j \leq m$ are given by

$$(A_n)_{ij} = \begin{cases} \lambda & \text{if } i = j = m, \\ 0 & \text{otherwise}, \end{cases}$$

$$(B_n)_{ij} = \begin{cases} -(\lambda + n\nu + i\mu) & \text{if } j = i,\ 0 \le i \le m-1, \\ -(\lambda + m\mu) & \text{if } j = i = m, \\ \lambda & \text{if } j = i+1,\ 0 \le i \le m-1, \\ i\mu & \text{if } j = i-1,\ 1 \le i \le m, \\ 0 & \text{otherwise}, \end{cases}$$

$$(C_n)_{ij} = \begin{cases} n\nu & \text{if } j = i+1,\ 0 \le i \le m-1, \\ 0 & \text{otherwise}. \end{cases}$$

Recall that N is the number of customers in the orbit at steady state and S is the number of busy servers at steady state. We denote the vector by

$$\boldsymbol{p}_n \equiv (p_{n0}, p_{n1}, \ldots, p_{nm}), \quad n = 0, 1, \ldots,$$

where $p_{ni} = \mathbb{P}(N = n, S = i),\ i = 0, 1, \ldots, m$, i.e., p_{ni} is the joint distribution of the number of customers in the orbit and the number of busy servers in the steady state.

The main interest of this paper is to investigate the behavior of $\boldsymbol{p}_n$ as n tends to infinity. Throughout the paper, for two sequences of real numbers $\{f_n : n = 0, 1, 2, \ldots\}$ and $\{g_n : n = 0, 1, 2, \ldots\}$, $f_n \sim g_n$ as $n \to \infty$ denotes $\lim_{n\to\infty} \frac{f_n}{g_n} = 1$. If $\{\boldsymbol{f}_n : n = 0, 1, 2, \ldots\}$ is a sequence of real vectors, and if $\{g_n : n = 0, 1, 2, \ldots\}$ is a sequence of real numbers, then

$$\boldsymbol{f}_n \sim g_n\, \boldsymbol{d} \quad \text{as } n \to \infty$$

for a vector $\boldsymbol{d}$, denotes $\lim_{n\to\infty} \frac{1}{g_n}\boldsymbol{f}_n = \boldsymbol{d}$.

3. Tail asymptotics of the queue size distribution

3.1. A general result

The following theorem presents the asymptotic behaviors of $p_{ni},\ i = 0, 1, \ldots, m$, as n tends to infinity. Because the proof is long and technical, it is deferred to Section 4.

Theorem 1. *For $i = 0, 1, \ldots, m$,*

$$p_{ni} \sim \frac{c}{i!}\left(\frac{\nu}{\mu}\right)^i n^{\frac{\lambda}{m\nu} - m + i} \rho^n \quad \text{as } n \to \infty, \tag{1}$$

where c is a positive constant.

In Section 4, it is proved that the constant c is represented as

$$c = \frac{(m-1)!\,(1-\rho)^{\frac{\lambda}{m\nu}}}{\Gamma\left(\frac{\lambda}{m\nu}\right)} \left(\frac{\mu}{\nu}\right)^{m-1} \mathbb{E}[(m-S)\rho^{m-S-1}] \exp\left(\int_1^{\frac{1}{\rho}} \frac{\mathbb{E}[h(z,S)z^N]}{\mathbb{E}[(m-S)\rho^{m-S-1}z^N]} dz\right), \tag{2}$$

where

$$h(z,i) = \frac{\lambda}{m\nu} \sum_{j=0}^{m-i-1} (\rho z)^j \frac{(j+1)(i - m\rho z)(1-z) + \rho(m-i)z}{\rho z^{m-i+1}} + \frac{\lambda}{m\nu}\left(\rho^{m-i-1}\frac{(i-mz)(m-i)+i}{z} - \frac{i}{\rho z^{m-i+1}}\right). \tag{3}$$

For the integral $\int_1^{\frac{1}{\rho}} \frac{\mathbb{E}[h(z,S)z^N]}{\mathbb{E}[(m-S)\rho^{m-S-1}z^N]} dz$ in (2), we need to check that (i) $0 < \mathbb{E}[(m-S)\rho^{m-S-1}z^N] < \infty$ for $z \in [1, \frac{1}{\rho})$; (ii) $h(z,S)z^N$ is integrable, i.e., $\mathbb{E}[|h(z,S)z^N|] < \infty$ for $z \in [1, \frac{1}{\rho})$; (iii) $\frac{\mathbb{E}[h(z,S)z^N]}{\mathbb{E}[(m-S)\rho^{m-S-1}z^N]}$ is an integrable function of z on $[1, \frac{1}{\rho})$. We observe that

(a) $(m-S)\rho^{m-S-1}$ is a nonnegative and bounded random variable that is positive with a positive probability for $z \in [1, \frac{1}{\rho})$;
(b) $h(z,S)$ is a bounded random variable for $z \in [1, \frac{1}{\rho})$;
(c) z^N is an integrable random variable for $z \in [1, \frac{1}{\rho})$, by Theorem 1.

Therefore, we have (i) by (a) and (c), and (ii) by (b) and (c). Further, from $h(z,m) = 0$, we have that

(d) there exists a finite number K such that, for all $z \in [1, \frac{1}{\rho})$, $|h(z,S)| \le K(m-S)\rho^{m-S-1}$.

Now we have (iii) by (d). In fact, (d) implies that $\frac{\mathbb{E}[h(z,S)z^N]}{\mathbb{E}[(m-S)\rho^{m-S-1}z^N]}$ is a bounded function of z on $[1, \frac{1}{\rho})$.

To obtain an explicit expression for the constant c, we need to know the generating functions,

$$\mathbb{E}[z^N \mathbb{1}_{\{S=i\}}] = \sum_{n=0}^{\infty} p_{ni} z^n, \quad i = 0, 1, \ldots, m-1. \tag{4}$$

However, there is no known practical exact formula for obtaining the generating functions. As shown in Section 5, we will use an accurate approximation method for obtaining the generating functions. As shown below, we will provide an explicit formula for the constant c that does not include any unknown factors, when $m = 1$ or 2.

3.2. *The cases of $m \le 2$*

In this subsection, we obtain explicit formulas for the constant c in the cases of $m \le 2$.

3.2.1. *The M/M/1 retrial queue*

When $m = 1$,

$$\mathbb{E}[(m-S)\rho^{m-S-1}] = \mathbb{P}(S = 0) = 1 - \rho. \tag{5}$$

By substituting $m = 1$ into (3), we have $h(z, 0) = h(z, 1) = 0$. Therefore, we have

$$\exp\left(\int_1^{\frac{1}{\rho}} \frac{\mathbb{E}[h(z,S)z^N]}{\mathbb{E}[(m-S)\rho^{m-S-1}z^N]} dz\right) = 1. \tag{6}$$

Substituting (5) and (6) into (2) yields $c = \frac{(1-\rho)^{\frac{\lambda}{\nu}+1}}{\Gamma\left(\frac{\lambda}{\nu}\right)}$. Therefore, for the M/M/1 retrial queue, Theorem 1 reduces to the following result, which is identical to the result of [17].

Theorem 2 ([17]). *For the* M/M/1 *retrial queue,*

$$p_{n0} \sim \frac{(1-\rho)^{\frac{\lambda}{\nu}+1}}{\Gamma\left(\frac{\lambda}{\nu}\right)} n^{\frac{\lambda}{\nu}-1} \rho^n \quad \text{as } n \to \infty,$$

$$p_{n1} \sim \frac{(1-\rho)^{\frac{\lambda}{\nu}+1}}{\Gamma\left(\frac{\lambda}{\nu}\right)} \frac{\nu}{\mu} n^{\frac{\lambda}{\nu}} \rho^n \quad \text{as } n \to \infty.$$

3.2.2. *The M/M/2 retrial queue*

When $m \ge 2$, it is difficult to express the constant c in (1) explicitly without unknown factors, because there is no known expression for the generating functions in (4). However, in the case of M/M/2 retrial queue, we can obtain the constant c explicitly using the results of Hanschke [15]. Hanschke [15] showed that

$$p_{n0} = c_4 \frac{(c_1)_n (c_2)_n}{(c_3)_n} \frac{\rho^n}{n!}, \tag{7}$$

$$p_{n1} = c_4 \frac{\nu}{\mu} \left(n + \frac{\lambda}{\nu}\right) \frac{(c_1)_n (c_2)_n}{(c_3)_n} \frac{\rho^n}{n!}, \tag{8}$$

$$p_{n2} = c_4 \frac{\nu^2}{2\mu^2} \frac{c_1 c_2}{c_3} \left(n + \frac{\lambda}{\nu} + \frac{\mu}{\nu} + 1\right) \frac{(c_1+1)_n (c_2+1)_n}{(c_3+1)_n} \frac{\rho^n}{n!}, \tag{9}$$

where

$$c_1 = \frac{2\lambda + \mu + \sqrt{4\lambda\mu + \mu^2}}{2\nu},$$

$$c_2 = \frac{2\lambda + \mu - \sqrt{4\lambda\mu + \mu^2}}{2\nu},$$

$$c_3 = \frac{3}{2}\frac{\lambda}{\nu} + \frac{\mu}{\nu} + 1,$$

$$c_4 = \frac{2\mu - \lambda}{\mu} \left[\frac{2\mu + \lambda}{\mu} \sum_{n=0}^{\infty} \frac{(c_1)_n (c_2)_n}{(c_3)_n} \frac{\rho^n}{n!} + \frac{\left(\frac{\lambda}{\mu}\right)^3}{3\lambda/\mu + 2\nu/\mu + 2} \sum_{n=0}^{\infty} \frac{(c_1+1)_n (c_2+1)_n}{(c_3+1)_n} \frac{\rho^n}{n!}\right]^{-1},$$

and

$$(x)_n = \begin{cases} 1 & \text{for } n = 0, \\ x(x+1)\cdots(x+n-1) & \text{for } n \geq 1. \end{cases}$$

A useful expression for the gamma function is given by (see p. 809 in [21])

$$\Gamma(x) = \lim_{n\to\infty} \frac{(n-1)!n^x}{x(x+1)\cdots(x+n-1)},$$

and so

$$(x)_n = x(x+1)\cdots(x+n-1) \sim \frac{(n-1)!n^x}{\Gamma(x)} \quad \text{as } n \to \infty.$$

Therefore, (7)–(9) imply respectively

$$p_{n0} \sim c_4 \frac{\Gamma(c_3)}{\Gamma(c_1)\Gamma(c_2)} n^{c_1+c_2-c_3-1}\rho^n \quad \text{as } n \to \infty,$$
$$p_{n1} \sim c_4 \frac{\nu}{\mu} \frac{\Gamma(c_3)}{\Gamma(c_1)\Gamma(c_2)} n^{c_1+c_2-c_3}\rho^n \quad \text{as } n \to \infty,$$
$$p_{n2} \sim c_4 \frac{\nu^2}{2\mu^2} \frac{\Gamma(c_3)}{\Gamma(c_1)\Gamma(c_2)} n^{c_1+c_2-c_3+1}\rho^n \quad \text{as } n \to \infty.$$

Thus we have the following theorem.

Theorem 3. *For the* M/M/2 *retrial queue,*

$$p_{n0} \sim c\, n^{\frac{\lambda}{2\nu}-2}\rho^n \quad \text{as } n \to \infty,$$
$$p_{n1} \sim c \frac{\nu}{\mu} n^{\frac{\lambda}{2\nu}-1}\rho^n \quad \text{as } n \to \infty,$$
$$p_{n2} \sim \frac{c}{2}\left(\frac{\nu}{\mu}\right)^2 n^{\frac{\lambda}{2\nu}}\rho^n \quad \text{as } n \to \infty,$$

where $c = \frac{c_4\Gamma(c_3)}{\Gamma(c_1)\Gamma(c_2)}$.

4. Proof of Theorem 1

In this section we prove Theorem 1. For this, we first consider the censored Markov process obtained by observing the process $N(t)$ and $S(t)$ only when the number of busy servers is less than or equal to $m-1$. The censored Markov process, denoted by $\{(\tilde{N}(t), \tilde{S}(t)) : t \geq 0\}$, is formally defined as

$$\tilde{N}(t) = N(\tau(t)), \qquad \tilde{S}(t) = S(\tau(t)),$$

where

$$\tau(t) \equiv \inf\left\{ s \geq 0 : \int_0^s \mathbb{1}_{\{S(u)\leq m-1\}} du = t \right\}.$$

The process $\{(\tilde{N}(t), \tilde{S}(t)) : t \geq 0\}$ is a Markov process with state space $\{0, 1, 2, \ldots\} \times \{0, 1, 2, \ldots, m-1\}$. The infinitesimal generator $\tilde{Q}$ of the Markov process $\{(\tilde{N}(t), \tilde{S}(t)) : t \geq 0\}$ is given by

$$\tilde{Q} = \begin{bmatrix} \tilde{A}_{00} & \tilde{A}_{01} & \tilde{A}_{02} & \tilde{A}_{03} & \cdots \\ \tilde{A}_{1,-1} & \tilde{A}_{10} & \tilde{A}_{11} & \tilde{A}_{12} & \cdots \\ O & \tilde{A}_{2,-1} & \tilde{A}_{20} & \tilde{A}_{21} & \cdots \\ O & O & \tilde{A}_{3,-1} & \tilde{A}_{30} & \cdots \\ & & & \ddots & \ddots \end{bmatrix},$$

where $\tilde{A}_{nk},\ k = -1, 0, 1, \ldots,$ are $m \times m$ matrices whose (i, j) components $(\tilde{A}_{nk})_{ij},\ 0 \leq i,\ j \leq m-1$ are given by

$$(\tilde{A}_{n,-1})_{ij} = \begin{cases} n\nu a_0 & \text{if } j = i = m-1, \\ n\nu & \text{if } j = i+1,\ 0 \leq i \leq m-2, \\ 0 & \text{otherwise,} \end{cases}$$

$$(\tilde{A}_{n0})_{ij} = \begin{cases} -(\lambda + n\nu + i\mu) & \text{if } j = i,\ 0 \le i \le m-2, \\ -(\lambda + n\nu + (m-1)\mu - \lambda a_0 - n\nu a_1) & \text{if } j = i = m-1, \\ \lambda & \text{if } j = i+1,\ 0 \le i \le m-2, \\ i\mu & \text{if } j = i-1,\ 1 \le i \le m-1, \\ 0 & \text{otherwise}, \end{cases}$$

$$(\tilde{A}_{nk})_{ij} = \begin{cases} \lambda a_k + n\nu a_{k+1} & \text{if } j = i = m-1, \\ 0 & \text{otherwise}, \end{cases} \quad k \ge 1,$$

with

$$a_n = \frac{1}{1+\rho}\left(\frac{\rho}{1+\rho}\right)^n, \quad n \ge 0.$$

Note that $a_n, n \ge 0$, is the probability that there are n exogenous arrivals during a period that starts when all servers are busy and ends when any one server completes its service.

By a well-known property of censored Markov processes, we have

$$\lim_{t\to\infty} \mathbb{P}(\tilde{N}(t) = n, \tilde{S}(t) = i) = \frac{1}{\sum_{k=0}^{\infty}\sum_{j=0}^{m-1} p_{kj}} p_{ni}, \quad n \ge 0,\ 0 \le i \le m-1.$$

We denote by

$$\tilde{\boldsymbol{p}}_n \equiv (p_{n0}, p_{n1}, \dots, p_{n,m-1})$$

and denote the vector generating function by

$$\tilde{\boldsymbol{P}}(z) \equiv (P_0(z), \dots, P_{m-1}(z)), \quad |z| \le 1,$$

where

$$P_i(z) \equiv \mathbb{E}[z^N \mathbb{1}_{\{S=i\}}] = \sum_{n=0}^{\infty} p_{ni} z^n, \quad 0 \le i \le m-1.$$

The vector generating function $\tilde{\boldsymbol{P}}(z)$ satisfies the following matrix differential equation. Because the proof is long and tedious, it is deferred to Appendix A. Here, $\tilde{\boldsymbol{P}}'(z)$ denotes the derivative of $\tilde{\boldsymbol{P}}(z)$ with respect to z and is interpreted componentwise.

Lemma 1. *For $z \in \mathbb{C}$ with $|z| \le 1$ and $z \ne 0$,*

$$\tilde{\boldsymbol{P}}'(z) = \tilde{\boldsymbol{P}}(z)\Big(\frac{\lambda}{m\nu}\frac{\rho}{1-\rho z}\boldsymbol{\xi}\boldsymbol{\eta} + \Psi(z)\Big),$$

where $\boldsymbol{\xi}$ is the m-dimensional column vector with ith component given by

$$\xi_i = (m-i)\rho^{m-i-1}, \quad 0 \le i \le m-1,$$

$\boldsymbol{\eta}$ is the m-dimensional row vector whose ith component is 0 if $0 \le i \le m-2$ and 1 if $i = m-1$, and $\Psi(z)$ is the $m \times m$ matrix with (i, j) component

$$(\Psi(z))_{ij} = \begin{cases} \dfrac{\lambda}{m\nu}\dfrac{i}{\rho z} & \text{if } j = i-1,\ 1 \le i \le m-1, \\ \dfrac{\lambda}{m\nu}\dfrac{i(1-z) - m\rho z}{\rho z^2} & \text{if } j = i,\ 0 \le i \le m-2, \\ \dfrac{\lambda}{m\nu}\dfrac{(i - m\rho z)(1-z)}{\rho z^{j-i+2}} & \text{if } i < j \le m-2,\ 0 \le i \le m-3, \\ \dfrac{\lambda}{m\nu}\left(\dfrac{m\rho z - i}{\rho z^{m-i}}\left(1 + \rho\displaystyle\sum_{k=0}^{m-i-1}(\rho z)^k\right) - m\rho^{m-i}\right) & \text{if } 0 \le i \le m-2, j = m-1, \\ -\dfrac{\lambda}{m\nu}\dfrac{(1+\rho)(m-1)}{\rho z} & \text{if } i = j = m-1, \\ 0 & \text{otherwise}. \end{cases} \tag{10}$$

We observe that all components of $\Psi(z)$ are rational functions of z that are analytic on $\mathbb{C} \setminus \{0\}$. By Lemma 1, we have the following corollary. The proof is given in Appendix B.

Corollary 1. *For $i = 0, 1, \ldots, m-1$, the functions $P_i(z)$, $|z| \le 1$, can be extended analytically to $\mathbb{C} \setminus [\frac{1}{\rho}, \infty)$.*

By Corollary 1, we see that if $|z| < \frac{1}{\rho}$, then $\mathbb{E}|z^N| < \infty$. Moreover, we know that the functions $P_i(z), 0 \le i \le m-1$, can be extended analytically to $\mathbb{C} \setminus [\frac{1}{\rho}, \infty)$. From now on, we use the same symbols $P_i(z),\ 0 \le i \le m-1$, to denote the extended functions. The vector valued function $\tilde{\boldsymbol{P}}(z) \equiv (P_0(z), \ldots, P_{m-1}(z)),\ z \in \mathbb{C} \setminus [\frac{1}{\rho}, \infty)$, has the extended functions $P_i(z), 0 \le i \le m-1$, as its components.

Lemma 2. *Suppose that $|z| < \frac{1}{\rho}$. For $1 \le z < \frac{1}{\rho}$, we have*

$$\tilde{\boldsymbol{P}}(z)\boldsymbol{\xi} = \mathbb{E}[(m-S)\rho^{m-S-1}z^N], \tag{11}$$

$$\tilde{\boldsymbol{P}}(z)\Psi(z)\boldsymbol{\xi} = \mathbb{E}[h(z,S)z^N]. \tag{12}$$

Proof. Suppose that $|z| < \frac{1}{\rho}$. First, Eq. (11) holds, because

$$\begin{aligned}\tilde{\boldsymbol{P}}(z)\boldsymbol{\xi} &= \sum_{i=0}^{m-1} P_i(z)\xi_i \\ &= \sum_{i=0}^{m-1} \mathbb{E}[z^N \mathbb{1}_{\{S=i\}}]\,(m-i)\rho^{m-i-1} \\ &= \mathbb{E}[(m-S)\rho^{m-S-1}z^N].\end{aligned}$$

Next we derive (12). By (10), we have for $0 \le i \le m-2$,

$$\begin{aligned}(\Psi(z)\boldsymbol{\xi})_i = {}& \frac{\lambda}{m\nu}\frac{i(m-i+1)\rho^{m-i}}{\rho z} + \frac{\lambda}{m\nu}\frac{(i(1-z)-m\rho z)(m-i)\rho^{m-i-1}}{\rho z^2} \\ &+ \sum_{j=i+1}^{m-2}\frac{\lambda}{m\nu}\frac{(i-m\rho z)(1-z)(m-j)\rho^{m-j-1}}{\rho z^{j-i+2}} + \frac{\lambda}{m\nu}\left(\frac{m\rho z-i}{\rho z^{m-i}}\left(1+\rho\sum_{k=0}^{m-i-1}(\rho z)^k\right) - m\rho^{m-i}\right),\end{aligned}$$

which can be written as the right-hand side of (3). By (10), we have for $i = m-1$,

$$(\Psi(z)\boldsymbol{\xi})_i = -\frac{\lambda}{m\nu}\frac{(1-\rho)(m-1)}{\rho z},$$

which is the same as the right-hand side of (3) for $i = m-1$. Therefore, we have $(\Psi(z)\boldsymbol{\xi})_i = h(z,i),\ 0 \le i \le m-1$. Thus

$$\begin{aligned}\tilde{\boldsymbol{P}}(z)\Psi(z)\boldsymbol{\xi} &= \sum_{i=0}^{m-1} P_i(z)h(z,i) \\ &= \sum_{i=0}^{m-1} \mathbb{E}[z^N \mathbb{1}_{\{S=i\}}]\,h(z,i).\end{aligned}$$

Since $h(z,m) = 0$, we have

$$\begin{aligned}\tilde{\boldsymbol{P}}(z)\Psi(z)\boldsymbol{\xi} &= \sum_{i=0}^{m} \mathbb{E}[z^N \mathbb{1}_{\{S=i\}}]\,h(z,i) \\ &= \mathbb{E}[h(z,S)z^N],\end{aligned}$$

which completes the proof of (12). $\square$

The following lemma plays a key role in proving Theorem 1. This lemma says that the formula (1) holds for the case of $i = m-1$. Here and subsequently, if a function has a removable singularity, then the value of the function at the singularity is interpreted by continuity.

Lemma 3. *We have*

$$p_{n,m-1} \sim \frac{c}{(m-1)!}\left(\frac{\nu}{\mu}\right)^{m-1} n^{\frac{\lambda}{m\nu}-1}\rho^n \quad \text{as } n \to \infty,$$

where c is given by (2).

Proof. According to Lemma 1 and Corollary 1,

$$\tilde{\boldsymbol{P}}'(z) = \frac{\lambda}{m\nu}\frac{\rho}{1-\rho z}\tilde{\boldsymbol{P}}(z)\boldsymbol{\xi}\boldsymbol{\eta} + \tilde{\boldsymbol{P}}(z)\Psi(z), \quad z \in \mathbb{C} \setminus \left[\frac{1}{\rho}, \infty\right). \tag{13}$$

Postmultiplying (13) by $\boldsymbol{\xi}$, we have

$$\tilde{\boldsymbol{P}}'(z)\boldsymbol{\xi} = \frac{\lambda}{m\nu}\frac{\rho}{1-\rho z}\tilde{\boldsymbol{P}}(z)\boldsymbol{\xi} + \tilde{\boldsymbol{P}}(z)\Psi(z)\boldsymbol{\xi}, \quad 1 \le z < \frac{1}{\rho}. \tag{14}$$

Letting $g(z) \equiv \frac{\tilde{\boldsymbol{P}}(z)\Psi(z)\boldsymbol{\xi}}{\tilde{\boldsymbol{P}}(z)\boldsymbol{\xi}}$, $1 \le z < \frac{1}{\rho}$, we have by Lemma 2

$$g(z) = \frac{\mathbb{E}[h(z,S)z^N]}{\mathbb{E}[(m-S)\rho^{m-S-1}z^N]}, \quad 1 \le z < \frac{1}{\rho} \tag{15}$$

and recall that the right-hand side of (15) is bounded in $z \in [1, \frac{1}{\rho})$. Eq. (14) is written as

$$\tilde{\boldsymbol{P}}'(z)\boldsymbol{\xi} = \tilde{\boldsymbol{P}}(z)\boldsymbol{\xi}\left[\frac{\lambda}{m\nu}\frac{\rho}{1-\rho z} + g(z)\right], \quad 1 \le z < \frac{1}{\rho}.$$

Solving the above differential equation gives

$$\tilde{\boldsymbol{P}}(z)\boldsymbol{\xi} = \tilde{\boldsymbol{P}}(1)\boldsymbol{\xi}\ \exp\left(\int_1^z g(u)du\right)\left(\frac{1-\rho z}{1-\rho}\right)^{-\frac{\lambda}{m\nu}}, \quad 1 \le z < \frac{1}{\rho},$$

which implies that

$$\tilde{\boldsymbol{P}}(z)\boldsymbol{\xi} \sim \tilde{\boldsymbol{P}}(1)\boldsymbol{\xi}\ \exp\left(\int_1^{\frac{1}{\rho}} g(u)du\right)\left(\frac{1-\rho z}{1-\rho}\right)^{-\frac{\lambda}{m\nu}} \quad \text{as } z \to \frac{1}{\rho}\text{- through the real line.} \tag{16}$$

By (13) and (14), we have

$$\lim_{z\to\frac{1}{\rho}-}\frac{P_i(z)}{\tilde{\boldsymbol{P}}(z)\boldsymbol{\xi}} = \lim_{z\to\frac{1}{\rho}-}\frac{P_i'(z)}{\tilde{\boldsymbol{P}}'(z)\boldsymbol{\xi}} = \eta_i, \tag{17}$$

where η_i, $0 \le i \le m-1$, is the ith component of the vector $\boldsymbol{\eta}$. According to (16) and (17),

$$\tilde{\boldsymbol{P}}(z) \sim \tilde{\boldsymbol{P}}(1)\boldsymbol{\xi}\ \exp\left(\int_1^{\frac{1}{\rho}} g(u)du\right)\left(\frac{1-\rho z}{1-\rho}\right)^{-\frac{\lambda}{m\nu}}\boldsymbol{\eta} \quad \text{as } z \to \frac{1}{\rho}\text{- through the real line.}$$

Therefore,

$$\tilde{\boldsymbol{P}}\left(\frac{\omega}{\rho}\right) \sim \tilde{\boldsymbol{P}}(1)\boldsymbol{\xi}\exp\left(\int_1^{\frac{1}{\rho}} g(u)du\right)(1-\rho)^{\frac{\lambda}{m\nu}}(1-\omega)^{-\frac{\lambda}{m\nu}}\boldsymbol{\eta} \quad \text{as } \omega \to 1\text{- through the real line,}$$

which can be written as

$$\tilde{\boldsymbol{P}}\left(\frac{\omega}{\rho}\right) \sim \frac{c}{(m-1)!}\left(\frac{\nu}{\mu}\right)^{m-1}\Gamma\left(\frac{\lambda}{m\nu}\right)(1-\omega)^{-\frac{\lambda}{m\nu}}\boldsymbol{\eta} \quad \text{as } \omega \to 1\text{- through the real line,} \tag{18}$$

where

$$c = \frac{(m-1)!(1-\rho)^{\frac{\lambda}{m\nu}}}{\Gamma\left(\frac{\lambda}{m\nu}\right)}\left(\frac{\mu}{\nu}\right)^{m-1}\tilde{\boldsymbol{P}}(1)\boldsymbol{\xi}\exp\left(\int_1^{\frac{1}{\rho}} g(z)dz\right). \tag{19}$$

Note that by Lemma 2, $\tilde{\boldsymbol{P}}(1)\boldsymbol{\xi} = \mathbb{E}[(m-S)\rho^{m-S-1}]$. Substituting this and (15) into (19), we have (2).

Now, the Karamata Tauberian theorem for power series (see Corollary 1.7.3 in [22]), applied to (18), gives

$$\sum_{k=0}^{n}\frac{1}{\rho^k}\tilde{\boldsymbol{p}}_k \sim \frac{c}{(m-1)!}\left(\frac{\nu}{\mu}\right)^{m-1}\frac{m\nu}{\lambda}n^{\frac{\lambda}{m\nu}}\boldsymbol{\eta} \quad \text{as } n \to \infty, \tag{20}$$

which implies

$$\frac{1}{\rho^n}\tilde{\boldsymbol{p}}_n = o(n^{\frac{\lambda}{m\nu}}) \quad \text{as } n \to \infty. \tag{21}$$

According to (13),

$$\tilde{\boldsymbol{P}}'\left(\frac{\omega}{\rho}\right) = \frac{\lambda}{m\nu}\frac{\rho}{1-\omega}\tilde{\boldsymbol{P}}\left(\frac{\omega}{\rho}\right)\boldsymbol{\xi\eta} + \tilde{\boldsymbol{P}}\left(\frac{\omega}{\rho}\right)\Psi\left(\frac{\omega}{\rho}\right), \quad \omega \in \mathbb{C}\setminus[1,\infty). \tag{22}$$

The arguments of Appendix B in [18], applied to (22), show that there is $\delta > 0$ satisfying

$$\tilde{\boldsymbol{P}}\left(\frac{\omega}{\rho}\right) = O\left(|1-\omega|^{-\delta}\right) \quad \text{as } \omega \to 1,\ |\omega| \le 1. \tag{23}$$

Let

$$\Psi_k = \frac{(-1)^k}{k!}\frac{d^k}{d\omega^k}\Psi\left(\frac{\omega}{\rho}\right)\bigg|_{\omega=1}, \quad k = 0, 1, \ldots, [\delta],$$

where $[\delta]$ stands for integer part of δ. Eq. (22) is written as

$$\tilde{\boldsymbol{P}}'\left(\frac{\omega}{\rho}\right) = \tilde{\boldsymbol{P}}\left(\frac{\omega}{\rho}\right)\left(\frac{\lambda}{m\nu}\frac{\rho}{1-\omega}\boldsymbol{\xi\eta} + \sum_{k=0}^{[\delta]}\Psi_k(1-\omega)^k\right) + \tilde{\boldsymbol{P}}\left(\frac{\omega}{\rho}\right)\widetilde{\Psi}(\omega), \quad \omega \in \mathbb{C}\setminus[1,\infty), \tag{24}$$

where $\widetilde{\Psi}(\omega) = \Psi(\frac{\omega}{\rho}) - \sum_{k=0}^{[\delta]}\Psi_k(1-\omega)^k$. Since $\tilde{\boldsymbol{P}}(\frac{\omega}{\rho})\Psi(\frac{\omega}{\rho})$ is analytic in ω on $\mathbb{C}\setminus[1,\infty)$ and $\widetilde{\Psi}(\omega) = O(|1-\omega|^{[\delta]+1})$ as $\omega \to 1, |\omega| \le 1$, (23) implies that $\tilde{\boldsymbol{P}}(\frac{\omega}{\rho})\widetilde{\Psi}(\omega)$ has a continuous extension to $\{\omega \in \mathbb{C} : |\omega| \le 1\}$. Therefore letting $\tilde{\boldsymbol{P}}\left(\frac{\omega}{\rho}\right)\widetilde{\Psi}(\omega) = \sum_{n=0}^{\infty}\psi_n\omega^n$ be the power series expansion, we have by the Riemann–Lebesgue lemma that

$$\lim_{n\to\infty}\psi_n = 0. \tag{25}$$

By comparing the coefficients of power series for both sides of (24), we see that

$$\frac{n+1}{\rho^n}\tilde{\boldsymbol{p}}_{n+1} = \frac{\lambda}{m\nu}\rho\sum_{k=0}^{n}\frac{1}{\rho^k}\tilde{\boldsymbol{p}}_k\boldsymbol{\xi\eta} + \sum_{k=0}^{[\delta]}\sum_{l=0}^{k}\binom{k}{l}\frac{1}{\rho^{n-l}}\tilde{\boldsymbol{p}}_{n-l}\Psi_k(-1)^l + \psi_n, \quad n \ge [\delta].$$

This equation together with (20), (21) and (25), gives

$$\frac{n+1}{\rho^n}\tilde{\boldsymbol{p}}_{n+1} \sim \frac{c}{(m-1)!}\left(\frac{\nu}{\mu}\right)^{m-1} n^{\frac{\lambda}{m\nu}}\rho\,\boldsymbol{\eta} \quad \text{as } n \to \infty,$$

which is written as

$$\tilde{\boldsymbol{p}}_n \sim \frac{c}{(m-1)!}\left(\frac{\nu}{\mu}\right)^{m-1} n^{\frac{\lambda}{m\nu}-1}\rho^n\boldsymbol{\eta} \quad \text{as } n \to \infty.$$

Since $\boldsymbol{\eta} = (0, 0, \ldots, 1)$ and $\tilde{\boldsymbol{p}}_n = (p_{n0}, p_{n1}, \ldots, p_{n,m-1})$, we have

$$p_{n,m-1} \sim \frac{c}{(m-1)!}\left(\frac{\nu}{\mu}\right)^{m-1} n^{\frac{\lambda}{m\nu}-1}\rho^n \quad \text{as } n \to \infty,$$

which completes the proof. □

Before proceeding, we note that the vector $\boldsymbol{p}_n = (p_{n0}, p_{n1}, \ldots, p_{nm})$ satisfies the balance equation $(\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots)Q = \boldsymbol{0}$. From this we have

$$(\lambda + i\mu + n\nu)p_{ni} = \lambda p_{n,i-1} + (n+1)\nu p_{n+1,i-1} + (i+1)\mu p_{n,i+1}, \quad 0 \le i \le m-1,\ n \ge 0, \tag{26}$$

with the convention that $p_{n,-1} = 0$ for $n \ge 0$.

Finally, to complete the proof of Theorem 1 we need the following lemma. The proof is given in Appendix C.

Lemma 4. *For* $i = 0, 1, \ldots, m-2$,

$$\limsup_{n\to\infty}\frac{p_{ni}}{n^{\frac{\lambda}{m\nu}-m+i}\rho^n} < \infty. \tag{27}$$

Now we are ready to prove Theorem 1.

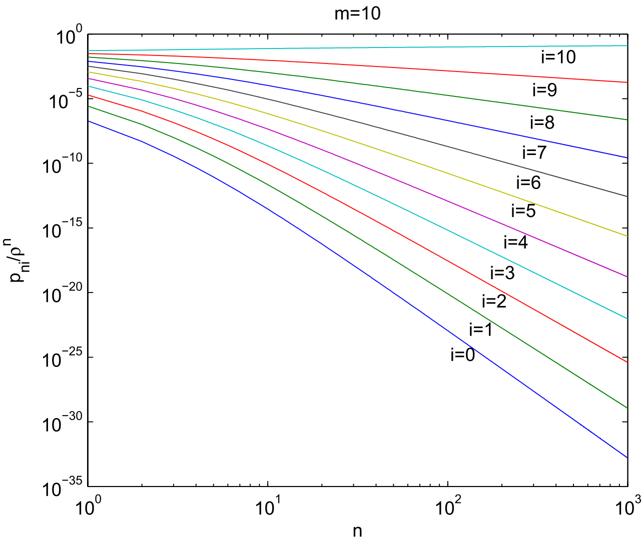

Fig. 1. Plots of $\frac{\mathbb{P}(N=n,S=i)}{\rho^n}$ for Example 1.

Proof of Theorem 1. Dividing (26) by $n^{\frac{\lambda}{m\nu}-m+i+1}\rho^n$ yields

$$\frac{(\lambda+i\mu+n\nu)p_{ni}}{n^{\frac{\lambda}{m\nu}-m+i+1}\rho^n}=\frac{\lambda p_{n,i-1}}{n^{\frac{\lambda}{m\nu}-m+i+1}\rho^n}+\frac{(n+1)\nu p_{n+1,i-1}}{n^{\frac{\lambda}{m\nu}-m+i+1}\rho^n}+\frac{(i+1)\mu p_{n,i+1}}{n^{\frac{\lambda}{m\nu}-m+i+1}\rho^n}$$

for $i=0,1,\ldots,m-1$. By Lemma 4, the first two terms on the right-hand side of the above equation go to 0 as $n\to\infty$, and thus we have

$$\lim_{n\to\infty}\frac{\nu p_{ni}}{n^{\frac{\lambda}{m\nu}-m+i}\rho^n}=\lim_{n\to\infty}\frac{(i+1)\mu\, p_{n,i+1}}{n^{\frac{\lambda}{m\nu}-m+i+1}\rho^n},\quad i=0,1,\ldots,m-1.$$

This and Lemma 3 complete the proof. □

5. Illustrations

In this section numerical examples are presented to illustrate our main result, Theorem 1. Theorem 1 asserts that for $i=0,1,\ldots,m$,

$$\frac{p_{ni}}{\rho^n}\sim\frac{c}{i!}\left(\frac{\nu}{\mu}\right)^i n^{\frac{\lambda}{m\nu}-m+i}\quad\text{as } n\to\infty. \tag{28}$$

From this we see that the exponent of the power function depends on the number of busy servers. Eq. (28) is equivalently expressed as

$$\log\frac{p_{ni}}{\rho^n}-\left(\frac{\lambda}{m\nu}-m+i\right)\log n\sim\log\left(\frac{c}{i!}\left(\frac{\nu}{\mu}\right)^i\right)\quad\text{as } n\to\infty. \tag{29}$$

Hence $\log\frac{p_{ni}}{\rho^n}$ becomes closer and closer to $\left(\frac{\lambda}{m\nu}-m+i\right)\log n+\log\left(\frac{c}{i!}\left(\frac{\nu}{\mu}\right)^i\right)$ as n tends to infinity. This point is demonstrated through numerical examples. We consider the following three models, all with retrial rate $\nu=1$.

Example 1 (*The* M/M/10 *Retrial Queue*)**.** We consider the M/M/10 retrial queue where the arrival rate is $\lambda=1$ and the mean service time is $\mu^{-1}=\frac{1}{7}$, and hence the offered load is $\rho=\frac{\lambda}{10\mu}=0.7$.

Example 2 (*The* M/M/50 *Retrial Queue*)**.** We consider the M/M/50 retrial queue where the arrival rate is $\lambda=1$ and the mean service time is $\mu^{-1}=\frac{1}{35}$, and hence the offered load is $\rho=0.7$.

Example 3 (*The* M/M/100 *Retrial Queue*)**.** We consider the M/M/100 retrial queue where the arrival rate is $\lambda=1$ and the mean service time is $\mu^{-1}=\frac{1}{70}$, and hence the offered load is $\rho=0.7$.

In Figs. 1, 3 and 5, for Examples 1–3, we plot $\frac{\mathbb{P}(N=n,S=i)}{\rho^n}$ versus n on a log–log scale. In the figures, the probabilities $p_{ni}=\mathbb{P}(N=n,S=i)$, $n=0,1,\ldots,i=0,\ldots,m$, are obtained as follows: It is known that the stationary queue

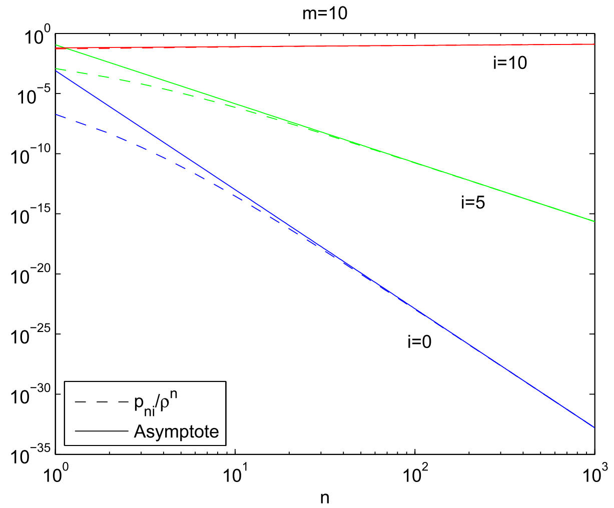

Fig. 2. Plots of $\frac{\mathbb{P}(N=n,S=i)}{\rho^n}$ and asymptotes for Example 1.

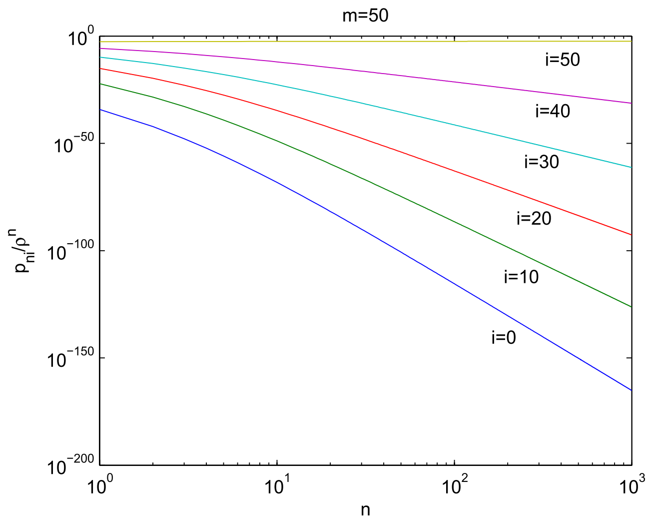

Fig. 3. Plots of $\frac{\mathbb{P}(N=n,S=i)}{\rho^n}$ for Example 2.

size of the M/M/m retrial queue with orbit of finite capacity K converges in distribution to that of the M/M/m retrial queue with infinite orbit capacity as K tends to infinity, i.e.,

$$\lim_{K\to\infty} p_{ni}^{(K)} = p_{ni}, \tag{30}$$

where $p_{ni}^{(K)}$ is the probability that there are n customers in the orbit and the number of busy servers is i at steady state in the M/M/m retrial queue with orbit of finite capacity K. Therefore, the probability p_{ni} is obtained as $p_{ni}^{(K)}$ such that $p_{ni}^{(K)}$ does not vary numerically as K increases. We observe that $\log \frac{\mathbb{P}(N=n,S=i)}{\rho^n}$ is asymptotically linear in $\log n$, as we expect from the formula (29).

In Figs. 2, 4 and 6, for Examples 1–3, we plot $\frac{\mathbb{P}(N=n,S=i)}{\rho^n}$ versus n on a log–log scale along with the values of the right-hand side of the formula (28) versus n. As demonstrated on the graphs for Figs. 2, 4 and 6, the right-hand side of (28) versus n are the asymptotic lines of the left-hand side of (28) versus n, on a log–log scale. To obtain the numeric values of the right-hand side of (28), we need to know the value c. The constant c is expressed in terms of the vector generating function $\tilde{\boldsymbol{P}}(z)$ according to (2) and Lemma 2. By numerically solving the differential equation (13) with the initial condition $\tilde{\boldsymbol{P}}(1) = (p_{00}, p_{01}, \ldots, p_{0,m-1})$, we can obtain $\tilde{\boldsymbol{P}}(z)$. Finally, in order to get the vector $\tilde{\boldsymbol{P}}(1) = (p_{00}, p_{01}, \ldots, p_{0,m-1})$, we use (30).

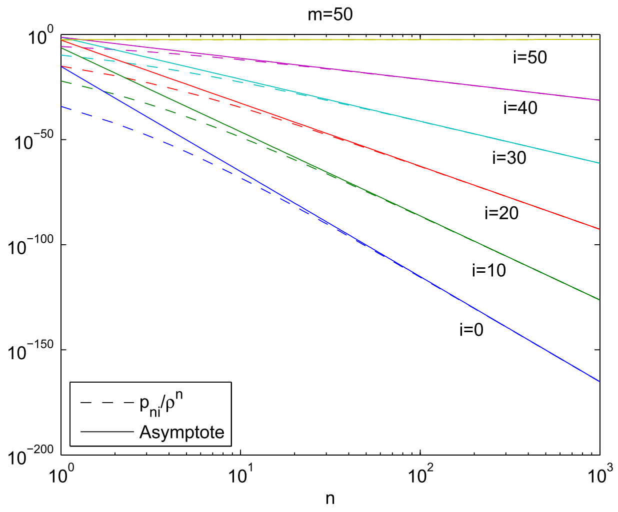

Fig. 4. Plots of $\frac{\mathbb{P}(N=n,S=i)}{\rho^n}$ and asymptotes for Example 2.

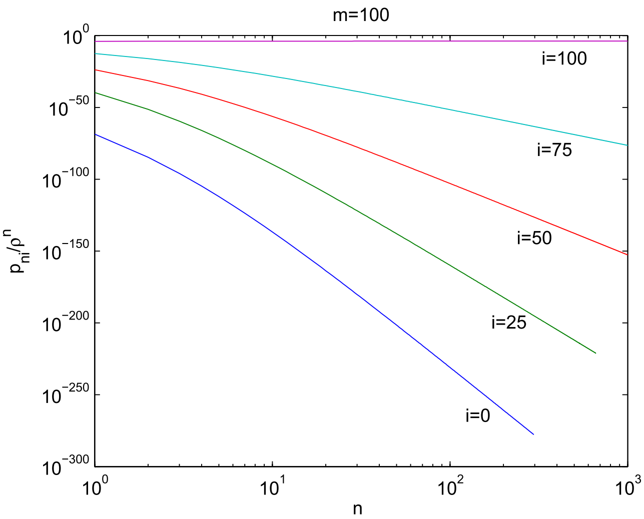

Fig. 5. Plots of $\frac{\mathbb{P}(N=n,S=i)}{\rho^n}$ for Example 3.

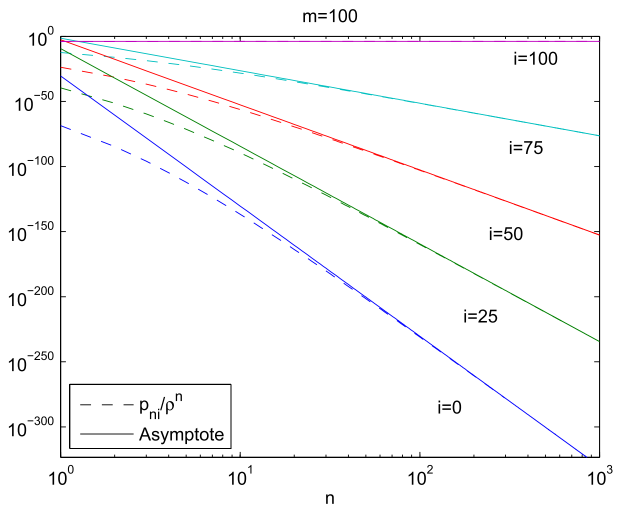

Fig. 6. Plots of $\frac{\mathbb{P}(N=n,S=i)}{\rho^n}$ and asymptotes for Example 3.

Acknowledgment

We would like to thank the anonymous referee for valuable comments which improved the original version.

Appendix A. Proof of Lemma 1

By the balance equation for the stationary distribution of the censored Markov process $\{(\tilde{N}(t), \tilde{S}(t)) : t \geq 0\}$, we have $\tilde{\boldsymbol{p}}\tilde{Q} = \mathbf{0}$, where $\tilde{\boldsymbol{p}} = (\tilde{\boldsymbol{p}}_0, \tilde{\boldsymbol{p}}_1, \tilde{\boldsymbol{p}}_2, \ldots)$. From this we get

$$\tilde{\boldsymbol{P}}'(z)B(z) = \tilde{\boldsymbol{P}}(z)C(z), \quad |z| \leq 1, \tag{31}$$

where $B(z)$ and $C(z)$ are given by

$$B(z) \equiv \nu(zI - \Phi(z)),$$
$$C(z) \equiv U - \lambda(I - \Phi(z)).$$

Here $\Phi(z)$ and U are $m \times m$ matrices whose (i, j) components are given by

$$(\Phi(z))_{ij} = \begin{cases} \dfrac{1}{1+\rho-\rho z} & \text{if } i = j = m-1, \\ 1 & \text{if } j = i+1,\ 0 \leq i \leq m-2, \\ 0 & \text{otherwise,} \end{cases}$$

$$U_{ij} = \begin{cases} -i\mu & \text{if } j = i,\ 0 \leq i \leq m-1, \\ i\mu & \text{if } j = i-1,\ 1 \leq i \leq m-1, \\ 0 & \text{otherwise.} \end{cases}$$

Eq. (31) is written as

$$\tilde{\boldsymbol{P}}'(z) = \tilde{\boldsymbol{P}}(z)C(z)(B(z))^{-1}, \tag{32}$$

for $z \in \{z \in \mathbb{C} : |z| \leq 1, \det(B(z)) \neq 0\}$. Since all components of $B(z)$ and $C(z)$ are rational functions of z, so are the components of $C(z)(B(z))^{-1}$.

Now we calculate the matrix $C(z)(B(z))^{-1}$ in (32). By direct calculation, we have

$$\begin{aligned} (B(z))^{-1} &= \frac{1}{\nu}\big(zI - \Phi(z)\big)^{-1} \\ &= \frac{1}{\nu}\frac{\rho z - \rho - 1}{(\rho z - 1)(z-1)}\boldsymbol{\zeta}(z)\boldsymbol{\eta} + \frac{1}{\nu}L(z), \end{aligned}$$

where $\boldsymbol{\zeta}(z)$ is the m-dimensional column vector whose jth component is given by

$$\zeta_j(z) = z^{j-(m-1)}, \quad 0 \leq j \leq m-1,$$

and $L(z)$ is the $m \times m$ matrix whose (i, j) component is given by

$$L_{ij}(z) = \begin{cases} z^{i-j-1} & \text{if } 0 \leq i \leq j \leq m-2, \\ 0 & \text{otherwise.} \end{cases}$$

Thus

$$\begin{aligned} C(z)\,(B(z))^{-1} &= \big(U - \lambda(I - \Phi(z))\big)\left(\frac{1}{\nu}\frac{\rho z - \rho - 1}{(\rho z - 1)(z-1)}\boldsymbol{\zeta}(z)\boldsymbol{\eta} + \frac{1}{\nu}L(z)\right) \\ &= \frac{1}{\nu}\frac{\rho z - \rho - 1}{(\rho z - 1)(z-1)}\big(U\boldsymbol{\zeta}(z) - \lambda(I - \Phi(z))\boldsymbol{\zeta}(z)\big)\boldsymbol{\eta} + \frac{1}{\nu}\big(UL(z) - \lambda(I - \Phi(z))L(z)\big). \end{aligned} \tag{33}$$

The jth components of $U\boldsymbol{\zeta}(z)$ and $(I - \Phi(z))\boldsymbol{\zeta}(z)$ are given respectively by

$$(U\boldsymbol{\zeta}(z))_j = j\mu(1-z)z^{j-m}, \quad 0 \leq j \leq m-1,$$

$$\big((I - \Phi(z))\boldsymbol{\zeta}(z)\big)_j = \begin{cases} (1-z)z^{j-(m-1)} & \text{if } 0 \leq j \leq m-2, \\ \dfrac{(1-z)\rho}{1+\rho-\rho z} & \text{if } j = m-1. \end{cases}$$

Hence the jth component of $U\boldsymbol{\zeta}(z) - \lambda(I - \Phi(z))\boldsymbol{\zeta}(z)$ is

$$\big(U\boldsymbol{\zeta}(z) - \lambda(I - \Phi(z))\boldsymbol{\zeta}(z)\big)_j = \begin{cases} (1-z)\left(\dfrac{j\mu}{z} - \lambda\right)z^{j-(m-1)} & \text{if } 0 \leq j \leq m-2, \\ (1-z)\left(\dfrac{(m-1)\mu}{z} - \dfrac{\lambda\rho}{1+\rho-\rho z}\right) & \text{if } j = m-1. \end{cases} \tag{34}$$

The (i,j) components of $UL(z)$ and $(I-\Phi(z))L(z)$ are given respectively by

$$(UL(z))_{ij} = \begin{cases} \dfrac{i\mu}{z} & \text{if } j = i-1,\ 1 \le i \le m-1, \\ i\mu\left(\dfrac{1}{z}-1\right)z^{i-j-1} & \text{if } 0 \le i \le j \le m-2, \\ 0 & \text{otherwise,} \end{cases}$$

$$\big((I-\Phi(z))L(z)\big)_{ij} = \begin{cases} \dfrac{1}{z} & \text{if } j = i,\ 0 \le i \le m-2, \\ \left(\dfrac{1}{z}-1\right)z^{i-j} & \text{if } 0 < j \le m-2,\ 0 \le i \le m-3, \\ 0 & \text{otherwise.} \end{cases}$$

Thus the (i,j) component of the second term on the right-hand side of (33) is

$$\frac{1}{\nu}\big(UL(z)-\lambda(I-\Phi(z))L(z)\big)_{ij} = \begin{cases} \dfrac{\lambda}{m\nu}\dfrac{i}{\rho z} & \text{if } j = i-1,\ 1 \le i \le m-1, \\ \dfrac{\lambda}{m\nu}\dfrac{i(1-z)-m\rho z}{\rho z^2} & \text{if } j = i,\ 0 \le i \le m-2, \\ \dfrac{\lambda}{m\nu}\dfrac{(i-m\rho z)(1-z)}{\rho z^{j-i+2}} & \text{if } i < j \le m-2,\ 0 \le i \le m-3, \\ 0 & \text{otherwise.} \end{cases} \tag{35}$$

By substituting (34) and (35) into (33), we obtain the (i,j) component of $C(z)(B(z))^{-1}$ as follows.

$$(C(z)(B(z))^{-1})_{ij} = \begin{cases} \dfrac{\lambda}{m\nu}\dfrac{i}{\rho z} & \text{if } j = i-1,\ 1 \le i \le m-1, \\ \dfrac{\lambda}{m\nu}\dfrac{i(1-z)-m\rho z}{\rho z^2} & \text{if } j = i,\ 0 \le i \le m-2, \\ \dfrac{\lambda}{m\nu}\dfrac{(i-m\rho z)(1-z)}{\rho z^{j-i+2}} & \text{if } i < j \le m-2,\ 0 \le i \le m-3, \\ \dfrac{\lambda}{m\nu}\left(\dfrac{\rho}{\rho z-1}-1\right)\left(\dfrac{i}{\rho z^{m-i}}-\dfrac{m}{z^{m-i-1}}\right) & \text{if } 0 \le i \le m-2, j = m-1, \\ \dfrac{\lambda}{m\nu}\left(\dfrac{\rho}{\rho z-1}-1\right)\dfrac{m-1}{\rho z}-\dfrac{\lambda}{m\nu}\dfrac{m\rho}{\rho z-1} & \text{if } i = j = m-1, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\boldsymbol{\xi} = (m\rho^{m-1}, (m-1)\rho^{m-2}, \ldots, 2\rho, 1)^\top$ and $\boldsymbol{\eta} = (0,\ldots,0,1)$, the (i,j) component of $C(z)(B(z))^{-1} - \frac{\lambda}{m\nu}\frac{\rho}{1-\rho z}\boldsymbol{\xi}\boldsymbol{\eta}$ is $(\Psi(z))_{ij}$ described in (10). That is,

$$C(z)(B(z))^{-1} = \frac{\lambda}{m\nu}\frac{\rho}{1-\rho z}\boldsymbol{\xi}\boldsymbol{\eta} + \Psi(z). \tag{36}$$

Finally, substituting (36) into (32) completes the proof of Lemma 1. □

Appendix B. Proof of Corollary 1

For $z \in \mathbb{C} \setminus ((-\infty,0] \cup [\frac{1}{\rho},\infty))$, consider the initial value problem:

$$\frac{d}{dt}\boldsymbol{x}(t) = (z-1)\boldsymbol{x}(t)\left(\frac{\lambda}{m\nu}\frac{\rho}{1-\rho(zt+1-t)}\boldsymbol{\xi}\boldsymbol{\eta} + \Psi(zt+1-t)\right), \quad 0 \le t \le 1,$$
$$\boldsymbol{x}(0) = \tilde{\boldsymbol{P}}(1).$$

Let $\boldsymbol{\alpha}(z,t)$ be the solution of the above initial value problem. Then we have the following properties:

(i) $\boldsymbol{\alpha}(z,1)$ is analytic in z on $\mathbb{C} \setminus ((-\infty,0] \cup [\frac{1}{\rho},\infty))$.

(ii) $\boldsymbol{\alpha}(z,1) = \tilde{\boldsymbol{P}}(z)$ for all complex numbers z with $|z| \le 1$ and $z \notin [-1,0]$.

For $z \in \mathbb{C} \setminus [0, \infty)$, consider the initial value problem:

$$\frac{d}{dt}\boldsymbol{x}(t) = (z+1)\boldsymbol{x}(t)\Big(\frac{\lambda}{m\nu}\frac{\rho}{1-\rho(zt-1+t)}\boldsymbol{\xi\eta} + \Psi(zt-1+t)\Big), \quad 0 \le t \le 1,$$
$$\boldsymbol{x}(0) = \tilde{\boldsymbol{P}}(-1).$$

Let $\boldsymbol{\beta}(z, t)$ be the solution of the above initial value problem. Then we have the following properties:

(iii) $\boldsymbol{\beta}(z, 1)$ is analytic in z on $\mathbb{C} \setminus [0, \infty)$.
(iv) $\boldsymbol{\beta}(z, 1) = \tilde{\boldsymbol{P}}(z)$ for all complex numbers z with $|z| \le 1$ and $z \notin [0, 1]$.

By (ii) and (iv) above, $\boldsymbol{\alpha}(z, 1) = \boldsymbol{\beta}(z, 1)$ for all complex numbers z with $|z| \le 1$ and $z \notin [-1, 1]$. This implies, by the identity theorem of analytic functions, that $\boldsymbol{\alpha}(z, 1) = \boldsymbol{\beta}(z, 1)$ for all $z \in \mathbb{C} \setminus \mathbb{R}$. Define $\boldsymbol{\gamma} : \mathbb{C} \setminus [\frac{1}{\rho}, \infty) \to \mathbb{C}^m$ as

$$\boldsymbol{\gamma}(z) = \begin{cases} \boldsymbol{\alpha}(z, 1) & \text{if } z \in \mathbb{C} \setminus \left((-\infty, 0] \cup \left[\dfrac{1}{\rho}, \infty\right)\right), \\ \boldsymbol{\beta}(z, 1) & \text{if } z \in \mathbb{C} \setminus [0, \infty), \\ \tilde{\boldsymbol{P}}(0) & \text{if } z = 0. \end{cases}$$

Note that $\boldsymbol{\gamma}$ is well defined. By (ii) and (iv) above, $\boldsymbol{\gamma}(z) = \tilde{\boldsymbol{P}}(z)$ for all complex numbers z with $|z| \le 1$. Hence $\boldsymbol{\gamma}(z)$ is analytic at $z = 0$. By (i) and (iii) above, $\boldsymbol{\gamma}(z)$ is also analytic on $\mathbb{C} \setminus (\{0\} \cup [\frac{1}{\rho}, \infty))$. Therefore $\boldsymbol{\gamma}(z)$ is analytic on $\mathbb{C} \setminus [\frac{1}{\rho}, \infty)$. Thus $\boldsymbol{\gamma}(z)$ is an analytic extension of $\tilde{\boldsymbol{P}}(z)$ to $\mathbb{C} \setminus [\frac{1}{\rho}, \infty)$, and the proof is complete. □

Appendix C. Proof of Lemma 4

To prove Lemma 4, we first show

$$\limsup_{n\to\infty} \frac{p_{ni}}{n^{\frac{\lambda}{m\nu}-1}\rho^n} < \infty, \quad i = 0, 1, \ldots, m-1. \tag{37}$$

This can be verified by reversed induction on i. By Lemma 3, (37) holds for $i = m-1$. Suppose that (37) holds for some $i \in \{1, 2, \ldots, m-1\}$. Then, by (26),

$$\begin{aligned} \limsup_{n\to\infty} \frac{p_{n,i-1}}{n^{\frac{\lambda}{m\nu}-1}\rho^n} &= \limsup_{n\to\infty} \frac{1}{\nu}\frac{(n+1)\nu p_{n+1,i-1}}{n^{\frac{\lambda}{m\nu}}\rho^n} \\ &\le \limsup_{n\to\infty} \frac{1}{\nu}\frac{(\lambda+i\mu+n\nu)p_{ni}}{n^{\frac{\lambda}{m\nu}}\rho^n} \\ &= \limsup_{n\to\infty} \frac{p_{ni}}{n^{\frac{\lambda}{m\nu}-1}\rho^n} \\ &< \infty, \end{aligned}$$

which means that (37) holds for $i-1$. Therefore, (37) holds for all $i = 0, 1, \ldots, m-1$.

Next we show

$$\limsup_{n\to\infty} \frac{p_{ni}}{n^{\frac{\lambda}{m\nu}-1+i}\rho^n} < \infty, \quad i = 0, 1, \ldots, m-1. \tag{38}$$

This can also be verified by induction on i. By (37), (38) holds for $i = 0$. Suppose that (38) holds for some $i \in \{0, 1, \ldots, m-2\}$. Then, by (26),

$$\begin{aligned} \limsup_{n\to\infty} \frac{p_{n,i+1}}{n^{\frac{\lambda}{m\nu}+i}\rho^n} &\le \limsup_{n\to\infty} \frac{\lambda+i\mu+n\nu}{(i+1)\mu}\frac{p_{ni}}{n^{\frac{\lambda}{m\nu}+i}\rho^n} \\ &= \limsup_{n\to\infty} \frac{\nu}{(i+1)\mu}\frac{p_{ni}}{n^{\frac{\lambda}{m\nu}-1+i}\rho^n} \\ &< \infty, \end{aligned}$$

which means that (38) holds for $i+1$. Therefore, (38) holds for all $i = 0, 1, \ldots, m-1$.

Finally, we prove (27) for all $i = 0, 1, \ldots, m-2$. Suppose that (27) does not hold for some $i \in \{0, 1, \ldots, m-2\}$, i.e.,

$$\limsup_{n\to\infty} \frac{p_{ni}}{n^{\frac{\lambda}{m\nu}-m+i}\rho^n} = \infty \quad \text{for some } i \in \{0, 1, \ldots, m-2\}. \tag{39}$$

Let $k_{\min}$ be the smallest positive integer k satisfying

$$\limsup_{n\to\infty} \frac{p_{ni}}{n^{\frac{\lambda}{m\nu}-k+i}\rho^n} = \infty \quad \text{for some } i \in \{0, 1, \ldots, m-2\}. \tag{40}$$

From (38) and (39), we have $2 \le k_{\min} \le m$. Let $i_{\max}$ be the largest integer $i \in \{0, 1, \ldots, m-2\}$ satisfying

$$\limsup_{n\to\infty} \frac{p_{ni}}{n^{\frac{\lambda}{m\nu}-k_{\min}+i}\rho^n} = \infty.$$

From (26) with $i = 0$, we have

$$\limsup_{n\to\infty} \frac{\lambda + n\nu}{n^{\frac{\lambda}{m\nu}-k_{\min}+1}\rho^n} p_{n0} = \limsup_{n\to\infty} \frac{\mu}{n^{\frac{\lambda}{m\nu}-k_{\min}+1}\rho^n} p_{n1}.$$

Therefore, if $\limsup_{n\to\infty} \frac{p_{n0}}{n^{\frac{\lambda}{m\nu}-k_{\min}}\rho^n} = \infty$, then $\limsup_{n\to\infty} \frac{p_{n1}}{n^{\frac{\lambda}{m\nu}-k_{\min}+1}\rho^n} = \infty$. Thus $1 \le i_{\max} \le m-2$. By (26) with $i = i_{\max}$,

$$\limsup_{n\to\infty} \frac{(\lambda + i_{\max}\mu + n\nu)p_{n,i_{\max}}}{n^{\frac{\lambda}{m\nu}-k_{\min}+i_{\max}+1}\rho^n} \le \limsup_{n\to\infty} \frac{\lambda p_{n,i_{\max}-1}}{n^{\frac{\lambda}{m\nu}-k_{\min}+i_{\max}+1}\rho^n} + \limsup_{n\to\infty} \frac{(n+1)\nu p_{n+1,i_{\max}-1}}{n^{\frac{\lambda}{m\nu}-k_{\min}+i_{\max}+1}\rho^n} + \limsup_{n\to\infty} \frac{(i_{\max}+1)\mu p_{n,i_{\max}+1}}{n^{\frac{\lambda}{m\nu}-k_{\min}+i_{\max}+1}\rho^n}.$$

Note that the first and third limits on the right-hand side of the above are finite by definitions of $k_{\min}$ and $i_{\max}$, respectively, whereas the limit of the left-hand side is infinite. Thus

$$\limsup_{n\to\infty} \frac{p_{n,i_{\max}-1}}{n^{\frac{\lambda}{m\nu}-k_{\min}+i_{\max}}\rho^n} = \infty.$$

Therefore (40) holds for $k = k_{\min} - 1$, which is a contradiction. This completes the proof of Lemma 4. □

References

[1] J.R. Artalejo, A classified bibliography of research on retrial queues: progress in 1990–1999, Top 7 (1999) 187–211.
[2] J.R. Artalejo, Accessible bibliography on retrial queues, Math. Comput. Modelling 30 (1999) 1–6.
[3] J.R. Artalejo, Accessible bibliography on retrial queues: progress in 2000–2009, Math. Comput. Modelling 51 (2010) 1071–1081.
[4] G.I. Falin, A survey of retrial queues, Queueing Syst. 7 (1990) 127–168.
[5] V.G. Kulkarni, H.M. Liang, Retrial queues revisited, in: J.H. Dshalalow (Ed.), Frontiers in Queueing: Models and Applications in Science and Engineering, CRC Press, Boca Raton, 1997, pp. 19–34.
[6] T. Yang, J.G.C. Templeton, A survey on retrial queues, Queueing Syst. 2 (1982) 201–233.
[7] J.R. Artalejo, A. Gómez-Corral, Retrial Queueing Systems, Springer, 2008.
[8] G.I. Falin, J.G.C. Templeton, Retrial Queues, Chapman & Hall, London, 1997.
[9] N. Deul, Stationary conditions for multiserver queueing systems with repeated calls, Elektron. Inf.verarb. Kybern. 16 (1980) 607–613.
[10] B.S. Greenberg, R.W. Wolff, An upper bound on the performance of queues with returning customers, J. Appl. Probab. 24 (1987) 466–475.
[11] C.E.M. Pearce, On the problem of re-attempted calls in teletraffic, Comm. Statist. Stochastic Models 3 (1987) 393–407.
[12] M.F. Neuts, B.M. Rao, Numerical investigation of a multiserver retrial model, Queueing Syst. 7 (1990) 169–190.
[13] R.I. Wilkinson, Theories for toll traffic engineering in the USA, Bell Syst. Tech. J. 35 (1956) 421–514.
[14] J. Keilson, J. Cozzolino, H. Young, A service system with unfilled requests repeated, Oper. Res. 16 (1968) 1126–1137.
[15] T. Hanschke, Explicit formulas for the characteristics of the M/M/2/2 queue with repeated attempts, J. Appl. Probab. 24 (1987) 486–494.
[16] R.D. Nobel, H.C. Tijms, Waiting-time probabilities in the M/G/1 retrial queue, Statist. Neerlandica 60 (2006) 73–78.
[17] J. Kim, B. Kim, S.-S. Ko, Tail asymptotics for the queue size distribution in an M/G/1 retrial queue, J. Appl. Probab. 44 (2007) 1111–1118.
[18] B. Kim, J. Kim, J. Kim, Tail asymptotics for the queue size distribution in the MAP/G/1 retrial queue, Queueing Syst. 66 (2010) 79–94.
[19] W. Shang, L. Liu, Q. Li, Tail asymptotics for the queue length in an M/G/1 retrial queue, Queueing Syst. 52 (2006) 193–198.
[20] J. Kim, J. Kim, B. Kim, Regularly varying tail of the waiting time distribution in M/G/1 retrial queue, Queueing Syst. 65 (2010) 365–383.
[21] A. Jeffrey, Linear Algebra and Ordinary Differential Equations, CRC Press, 1993.
[22] N.H. Bingham, C.M. Goldie, J.L. Teugels, Regular Variation, Cambridge University Press, 1987.